

(Appendix A)

Strausberg R.;
Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
EMBL: BC033877; AAH3877.1; -
Hypothetical protein.
SEQUENCE 605 AA; 1182534F9556DADD CRC64;

Query Match 100.0%; Score 2314; DB 4; Length 605;
Local Similarity 99.8%; Pred. No. 3.8e-181;
ches 444; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

1 MSGRDTLIGLCLALALSLAMMFFRITTLVHIFISLVILGLFVCGVLMWLYDYTN 60
161 MSGRDTLIGLCLALALSLAMMFFRITTLVHIFISLVILGLFVCGVLMWLYDYTN 220
61 DLSIEDTERENMKCVLGAIVSTGITAVALVLFVLRKRIKLTVELFQITNKAISSAPF 120
221 DLSIEDTERENMKCVLGAIVSTGITAVALVLFVLRKRIKLTVELFQITNKAISSAPF 280
121 LIFQPLMTFALLIFFWLWVAVLISLGTAGAAQVWEGGVVEKPLSGIRYMSYHLIGLI 180
281 LIFQPLMTFALLIFFWLWVAVLISLGTAGAAQVWEGGVVEKPLSGIRYMSYHLIGLI 340
181 WTSEFLLACQMTAGAVTTCYFNRSKNDPDPHPLISSLSLFFHOGTIVKGSFLISVY 240
341 WTSEFLLACQMTAGAVTTCYFNRSKNDPDPHPLISSLSLFFHOGTIVKGSFLISVY 400
241 RIPRIIWMQNALKEOQHGALSRYLFRCCYCCFWCLDKYLHINQNAVTTTAINGTDFC 300
401 RIPRIIWMQNALKEOQHGALSRYLFRCCYCCFWCLDKYLHINQNAVTTTAINGTDFC 460
301 TSADAFKILSKNSHFTSINCGRDFIIFLKGVLVVCFTVGGIMAFNNRPAQVAVPL 360
461 TSADAFKILSKNSHFTSINCGRDFIIFLKGVLVVCFTVGGIMAFNNRPAQVAVPL 520
361 LVAFAVYLVASHFLSVFETVDALFLCPAVDLETDGSSSEKPYFMDQELSPVKSINKL 420
521 LVAFAVYLVASHFLSVFETVDALFLCPAVDLETDGSSSEKPYFMDQELSPVKSINKL 580
421 NNARAQODKHSLSRNEGETELQAVR 445
581 NNARAQODKHSLSRNEGETELQAVR 605

```

UNT 2
1AM1 PRELIMINARY; PRT; 677 AA.
Q8N4M1;
01-OCT-2002 (TREMBlrel. 22, Created)
01-JUN-2003 (TREMBlrel. 24, Last sequence update)
01-OCT-2003 (TREMBlrel. 25, Last annotation update)
Hypothetical protein MG045474 (Fragment).
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
(1)
SEQUENCE FROM N.A.
Tissue=Brain, Lung, and Testis;
Strausberg R.;
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
EMBL: BC033858; AAH33858.2; -
InterPro: IPR007603; DUF580.
Pfam: PF04515; DUF580; 1.
Hypothetical protein.
NON TER
1
SEQUENCE 677 AA; 75876 MW; 115675C8C0C2D02E CRC64;

Query Match 100.0%; Score 2314; DB 4; Length 677;
Local Similarity 99.8%; Pred. No. 4.3e-181;
Mismatches 0; Indels 0; Gaps 0;

```

DB 233 MSGRDTLIGLCLALALSLAMMFFRITTLVHIFISLVILGLFVCGVLMWLYDYTN 292
QY 61 DLSIEDTERENMKCVLGAIVSTGITAVALVLFVLRKRIKLTVELFQITNKAISSAPF 120
DB 293 DLSIEDTERENMKCVLGAIVSTGITAVALVLFVLRKRIKLTVELFQITNKAISSAPF 352
QY 121 LIFQPLMTFALLIFFWLWVAVLISLGTAGAAQVWEGGVVEKPLSGIRYMSYHLIGLI 180
DB 353 LIFQPLMTFALLIFFWLWVAVLISLGTAGAAQVWEGGVVEKPLSGIRYMSYHLIGLI 412
QY 181 WTSEFLLACQMTAGAVTTCYFNRSKNDPDPHPLISSLSLFFHOGTIVKGSFLISVY 240
DB 413 WTSEFLLACQMTAGAVTTCYFNRSKNDPDPHPLISSLSLFFHOGTIVKGSFLISVY 472
QY 241 RIPRIIWMQNALKEOQHGALSRYLFRCCYCCFWCLDKYLHINQNAVTTTAINGTDFC 300
DB 473 RIPRIIWMQNALKEOQHGALSRYLFRCCYCCFWCLDKYLHINQNAVTTTAINGTDFC 532
QY 301 TSADAFKILSKNSHFTSINCGRDFIIFLKGVLVVCFTVGGIMAFNNRPAQVAVPL 360
DB 533 TSADAFKILSKNSHFTSINCGRDFIIFLKGVLVVCFTVGGIMAFNNRPAQVAVPL 592
QY 361 LVAFAVYLVASHFLSVFETVDALFLCPAVDLETDGSSSEKPYFMDQELSPVKSINKL 420
DB 593 LVAFAVYLVASHFLSVFETVDALFLCPAVDLETDGSSSEKPYFMDQELSPVKSINKL 652
QY 421 NNARAQODKHSLSRNEGETELQAVR 445
DB 653 NNARAQODKHSLSRNEGETELQAVR 677

```

RESULT 3
Q8BWY7 PRELIMINARY; PRT; 428 AA.
ID Q8BWY7;
AC Q8BWY7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE D639P13.1 (Novel protein similar to rat transporter-like protein
(CTR1)) (Fragment).
DN D639P13.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
(1)
SEQUENCE FROM N.A.
RP
RA Collier R.;
RU Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL359554; CAC36091.1; -
DR InterPro: IPR007603; DUF580.
DR Pfam: PF04515; DUF580; 1.
NON TER
1
SEQUENCE 428 AA; 46789 MW; 977F4793F72BA5A7 CRC64;

Query Match 95.9%; Score 2220.5; DB 4; Length 428;
Best Local Similarity 99.5%; Pred. No. 1.2e-173;
Matches 427; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```

QY 17 LSLAMMFFRITTLVHIFISLVILGLFVCGVLMWLYDYTNLSIEDTERENMKCV 76
DB 1 LSLAMMFFRITTLVHIFISLVILGLFVCGVLMWLYDYTNLSIEDTERENMKCV 60
QY 77 LGRVISTGTITAVLVLFVLRKRIKLTVELFQITNKAISSAPFLQPLMTFALLIFFW 136
DB 61 LGRVISTGTITAVLVLFVLRKRIKLTVELFQITNKAISSAPFLQPLMTFALLIFFW 120
QY 137 VLVAVVLSLGTAGAAQVWEGGVVEKPLSGIRYMSYHLIGLIWTSEFLLACQMTIAG 196
DB 121 VLVAVVLSLGTAGAAQVWEGGVVEKPLSGIRYMSYHLIGLIWTSEFLLACQMTIAG 180

```

Db 181 AVATCYENRSGNDPDPHPIISLSILFYHOGITVXGSLISVRIPIRIWMQNALKE 240
 QY 257 QOHGALSRYLFRCCYCCFWCLDKYLHLNQNAYTTAINGTDPCTSADAFKILSKNSH 316
 Db 241 -QHGAISRYLFRCCYCCFWCLDKYLHLNQNAYTTAINGTDPCTSADAFKILSKNSH 299
 QY 317 FTSINGEGDPIIFLGKLVVCFVFGGLAFVYNRAFOYMAVPLLVAFFALVAHSPLS 376
 Db 300 FTSINGEGDPIIFLGKLVVCFVFGGLAFVYNRAFOYMAVPLLVAFFALVAHSPLS 359
 QY 377 VFEETVLDALFLCPAVDLETNDGSEKPYFMDQEFISFVRSNKNLNARAAQDKSLRNE 436
 Db 360 VFEETVLDALFLCPAVDLETNDGSEKPYFMDQEFISFVRSNKNLNARAAQDKSLRNE 419
 QY 437 GTELOAIVR 445
 Db 420 GTELOAIVR 428

RESULT 4

Q921V7 PRELIMINARY; PRT; 605 AA.
 ID 0921V7
 AC 0921V7
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Similar to transporter-like protein (Hypothetical protein).
 GN BC010552.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC010552; AAH10552.1; -
 DR EMBL; BC025548; AAH25548.1; -
 DR MGD; MGI:2384860; BC010552.
 DR InterPro; IPR007603; DUF580.
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF04515; DUF580.1.
 DR PROSITE; PS00280; IG_MHC; 1.
 KW Hypothetical Protein.
 SQ SEQUENCE 605 AA; 67303 MW; 242C506A13089A5 CRC64;

Query Match 80.2%; Score 1857; DB 11; Length 605;
 Best Local Similarity 78.4%; Pred. No. 9, 6e-144;
 Matches 349; Conservative 39; Mismatches 57; Indels 0; Gaps 0;

QY 1 MSGRDTILGCIILALSLAMMFTFRITTLVHIFISLVILGLFVCGVLMWLYDYTN 60
 Db 161 MGRDITILGCIILALSLAMMFTFRITTLVHIFISLVILGLFVCGVLMWLYDYTN 220
 QY 61 DSIETDERENKCVLGFRAIVSTGITAVLVILFRLRRLCTVLELHVNKAISSPF 120
 Db 221 DSIETDERENKCVLGFRAIVSTGITAVLVILFRLRRLCTVLELHVNKAISSPF 280
 QY 121 LTFQPLMTFALLFFVWLVAVVLSLGTAGAAQVWEGQVEYKPLSGIRYMSYHLIGI 180
 Db 281 LTFQPLMTFALLFFVWLVAVVLSLGTAGAAQVWEGQVEYKPLSGIRYMSYHLIGI 340
 QY 181 WTSEFLAACQQTIAAVVTCYFNRSKNDPDPHPIISLSILFFHOGTVKGSFLISV 240
 Db 341 WTSEFLAACQQTIAAVVTCYFNRSKNDPDPHPIISLSILFFHOGTVKGSFLISV 400
 QY 241 RIPRIITVMONALKEQHGALSRYLFRCCYCCFWCLDKYLHLNQNAYTTAINGTDFC 300
 Db 401 RIPRIITVMONALKEQHGALSRYLFRCCYCCFWCLDKYLHLNQNAYTTAINGTDFC 460
 QY 301 TSAKDAFKILSKNSHFTSINGEGDPIIFLGKLVVCFVFGGLAFVYNRAFOYMAVPL 360
 Db 461 TSAKDAFKILSKNSHFTSINGEGDPIIFLGKLVVCFVFGGLAFVYNRAFOYMAVPL 520

QY 361 LVAFAFVAVHSHFSLVFETVDALFLCPAVDLETNDGSEKPYFMDQEFISFVRSNKL 420
 Db 521 LVAFAFVAVHSHFSLVFETVDALFLCPAVDLETNDGSEKPYFMDQEFISFVRSNKL 580
 QY 421 NNAAQODKSLNREGTELOAIVR 445
 Db 581 NNAAQODKSLNREGTELOAIVR 605

RESULT 5

Q9JUZ7 PRELIMINARY; PRT; 653 AA.
 ID 09JUZ7
 AC 09JUZ7
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
 DE CTL1 protein.
 GN CTL1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue=Brain;
 RX MEDLINE-20144127; PubMed=10677542;
 RA O'Regan S., Traiffort E., Rust M., Cha N., Compagno D., Meunier F.M.;
 RT "An electric lobe suppressor for a yeast choline transporter mutation
 belongs to a new family of transporter-like proteins."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1835-1840 (2000).
 DR EMBL; AJ245619; CAB75555.1; -
 DR InterPro; IPR007603; DUF580.
 DR Pfam; PF04515; DUF580.1.
 SQ SEQUENCE 653 AA; 73045 MW; 3750C455A98D2F9A CRC64;

Query Match 48.3%; Score 1132; DB 11; Length 653;
 Best Local Similarity 50.0%; Pred. No. 2, 3e-94;
 Matches 212; Conservative 81; Mismatches 125; Indels 6; Gaps 2;

QY 1 MSGRDTILGCIILALSLAMMFTFRITTLVHIFISLVILGLFVCGVLMWLYDYTN 60
 Db 208 MTSKRIILGCIILALSLAMMFTFRITTLVHIFISLVILGLFVCGVLMWLYDYTN 267
 QY 61 DSIETDERENKCVLGFRAIVSTGITAVLVILFRLRRLCTVLELHVNKAISSPF 115
 Db 268 SPEKTVIPBGLQIAEDNLRALLLYALSAVTFVILFLIMLVKRVALLTALPHVAGKVF 327
 QY 116 SSAPFLFQPLMTFALLFFVWLVAVVLSLGTAGAAQVWEGQVEYKPLSGIRYMSYH 175
 Db 328 IHPFLVQPFMTFALLFFVWLVAVVLSLGTAGAAQVWEGQVEYKPLSGIRYMSYH 387
 QY 176 LIGIHTSEFLAACQQTIAAVVTCYFNRSKNDPDPHPIISLSILFFHOGTVKGSFLISV 235
 Db 388 VGLIHTSEFLAACQQTIAAVVTCYFNRSKNDPDPHPIISLSILFFHOGTVKGSFLISV 447
 QY 236 LISVRIPIITVMONALKEQHGALSRYLFRCCYCCFWCLDKYLHLNQNAYTTAINGTDFC 295
 Db 448 IITVRIPIITVMONALKEQHGALSRYLFRCCYCCFWCLDKYLHLNQNAYTTAINGTDFC 506
 QY 296 GTPCCTSAKDAFKILSKNSHFTSINGEGDPIIFLGKLVVCFVFGGLAFVYNRAFOY 355
 Db 507 GTPCCTSAKDAFKILSKNSHFTSINGEGDPIIFLGKLVVCFVFGGLAFVYNRAFOY 566
 QY 356 WAVEPLLVAFVAVHSHFSLVFETVDALFLCPAVDLETNDGSEKPYFMDQEFISFVRSNKL 415
 Db 567 WAVEPLLVAFVAVHSHFSLVFETVDALFLCPAVDLETNDGSEKPYFMDQEFISFVRSNKL 626
 QY 416 NSRK 419
 Db 627 NSRK 630